Parametric Estimation of Crystallographic Texture Using Estimation Maximization

Carl Ahlborg

April 13, 2018

Materials Science and Engineering Department

Dr. Stephen R. Niezgoda

Abstract

The properties of materials are functions of their internal structure. Most engineering materials, including metals, are polycrystalline and composed of microscopic crystals. The crystallographic texture, the description of the orientations of the crystallites in a material, is a key structural indicator of the deformation behavior. Estimating the orientation distribution function (ODF) of a sample allows materials scientists to identify the probability a crystal in a sample is oriented in a certain direction. The current process, which employs a Fourier series expansion over a generalized spherical harmonic basis, leaves much to be desired and is poorly understood by much of the community that uses it. Some limitations include bias introduced by ad-hoc parameters and poor accuracy for small sample sets. The purpose of this study is to develop an algorithm to estimate the ODF using a mixture model that would be free from ad-hoc parameters. To accomplish this, the algorithm uses a mixture of symmetrized Bingham distributions. Using these distributions, it employs a Estimation Maximization (EM) approach to estimate the distribution and reevaluate the data to improve subsequent estimations. It also used a minimum message length (MML) criterion to prevent overfitting, or making too specific estimations based on insufficient data. This algorithm is compared with a similar algorithm developed in tandem using a mixture of symmetrized Bingham distributions but using a Markov Chain Monte Carlo (MCMC) approach instead.

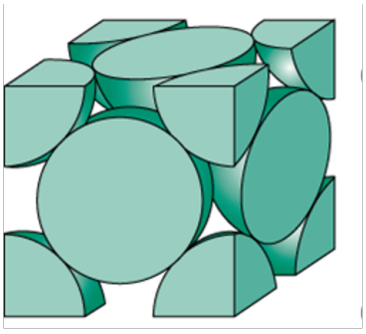
Aknowledgements

This research was funded as part of the Engineering Summer Research Scholarship Program, supported by The Ohio State University College of Engineering. I would like to thank my advisor, Dr. Stephen Niezgoda, for his guidance during the project, as well as his laboratory group, the Mesoscale Mechanics and Microstructures Laboratory, for their support and help. I would also like to thank James Matuk and Prof. Oksana Chkrebtii in the Statistics Department for their work on the MCMC algorithm.

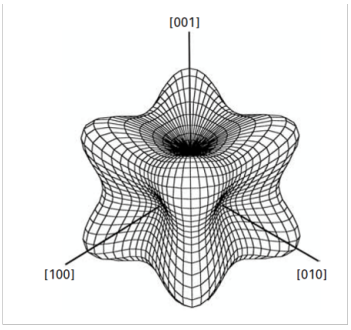
Contents

1 Introduction

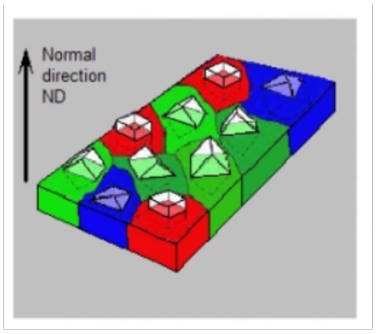
The properties of materials are functions of their internal structure. Most advanced structural materials, including metals, are polycrystalline and composed of millions of microscopic crystals. This internal structure evolves during processing and the processing history of a material can be as important as the chemical composition of an alloy for determining properties and performance. The crystallographic texture, or preferential orientation of the little crystals in a material, is a key structural descriptor for predicting the deformation behavior. The texture is typically characterized by the orientation distribution function (ODF) which describes the probability of a crystal in a sample is facing a certain direction, or equivalently the volume fraction of crystals with a given orientation in the samples.



FCC crystal
Retrieved from socratic.org



Elastic modulus along different orientations
M.A. Meyers, K.K. Chumar, Mechanical Behaviors of Materials



Polycrystalline sample, different orientations
Retrieved from ebsd.info

Such micrographs are routinely measured in both research and industrial settings. The current process for estimating the ODF from such maps employs a Fourier series expansion over a generalized spherical harmonic basis. However this procedure, while routine and largely automated, is really more qualitative or semi-quantitative and is not suited for quantitative comparison between samples or for comparison between experimental datasets and those produced by modeling and simulation. Some limitations include bias introduced by ad-hoc parameters, such as bandwidth of the Fourier

series and halfwidth of a smoothing kernel. Also, the method is known to perform poorly for small sample sizes. In contrast, modeling the ODF using a mixture of symmetrized Bingham distributions offers a method for quantitative comparison, bypassing many of the concerns listed above. This paper builds on previous work by Prof. Niezgoda in development of the symmetrized Bingham distrution mixture model and application of estimation maximization (EM) to Bingham distributions. The focus of this paper is on using estimation maximization (EM) to fit a symmetrized Bingham distribution mixture model to a set of sample orientations. The motivations behind developing this algorithm, over continued use of the Fourier series expansion method, include eliminating ad-hoc parameters, reducing overfitting on small samples, and allowing for easy calculation of uncertainty quantification.

2 Background

This research project built off the work of Prof. Niezgoda on crystallographic texture analysis. Together with Dr. Jared Glover (Computer Science and Artificial Intelligence Laboratory MIT), he demonstrated how unsupervised learning of mixture models can be applied to a 3-dimensional rotations. For this work they adopted the Bingham distribution. Recently, along with MSE BS and now Ph.D. student Eric Magnusion, they developed a symmetrized Bingham distribution, which extends the Bingham distribution to the space of crystal orientations. They also developed computationally efficient tools for the estimation of parameters for the symmetrized Bingham distribution[2]. This paper builds on these two research results and applies this new symmetrized Bingham distribution into the EM unsupervised learning approach developed by Prof. Niezgoda, to measure the ODF for real engineering materials. Another algorithm was developed at the same time

by James Matuk in the Statistics Department. While also using a mixture model of symmetrized Bingham distributions, Matuk's approach uses a Markov Chain Monte Carlo (MCMC) approach to fit the parameters for the distribution.

2.1 Quaternion orientation and crystal symmetry

Orientations are represented in a number of different ways, most popularly using Euler angles. In this paper, quaternion orientations are used instead. A quaternion is a four dimensional vector extended to the complex numbers, expressed in the form

$$a + bi + cj + dk$$

where a, b, c, and d are real numbers while i, j, and k are the fundamental quaternion units. Quaternions are multiplied using the table below.

Quaternion multiplication

×	**1**	***i***	***j***	***k***
1	1	*i*	*j*	*k*
i	*i*	−1	*k*	−*j*
j	*j*	−*k*	−1	*i*
k	*k*	*j*	−*i*	−1

Additionally, a quaternion orientation is a unit quaternion and must satisfy $a^2 + b^2 + c^2 + d^2 = 1$. Because they are unit vectors of length 4, quaternions occupy the surface of the 3-sphere ($\mathbb{S}^3$). Quaternion orientations are used in this paper because of existing computational tools and the ease with which they can be rotated using multiplication.

Different materials have different crystal structures. Because crystals often contain fields of symmetry, any distribution describing the ODF must treat orientations that are identical by rotation as the same.

2.2 Bingham distribution

Because quaternion orientations occupy the surface of a hypersphere $\mathbb{S}^3$ as opposed to $\mathbb{R}^n$, a distribution over that space must be selected. Just as the Gaussian distribution is frequently used on real-value data on $(-\infty, \infty)$ because it is the maximum entropy distribution, the Bingham distribution is the maximum entropy distribution of unit quaternions. The Bingham distribution is an antipodally symmetric distribution on the unit hypersphere $\mathbb{S}^d \in \mathbb{R}^{d+1}$ for d = 3 and d= 7. For an orientation $\mathbf{g}$ in unit quaternion form, the pdf is given by

$$p(\mathbf{g}; \mathbf{\Lambda}, \mathbf{V}) = \frac{1}{F(\mathbf{\Lambda})} \exp\left(\sum_{i=1}^{4} \lambda_i (\mathbf{v}_i^T \mathbf{g})^2 \right)$$

where $\mathbf{\Lambda}$ is a vector of concentration parameters λ_i, $\mathbf{V}$ is a matrix with 4 columns of orthogonal unit quaternions $\mathbf{v}_i$ representing the principapal directions of the distribution, and F is a normalization constant. The parameters $\mathbf{\Lambda}, \mathbf{V}$ can be estimated for a set of N discrete orientations, $\mathbf{G} = \{\mathbf{g}^{(1)}, ..., \mathbf{g}^{(N)}\}$. For the scatter matrix $S = 1/N \sum_i \mathbf{g}^{(i)} \mathbf{g}^{(i)^T} = E[\mathbf{g}\mathbf{g}^T]$, $\mathbf{\Lambda}$ is determined from the eigenvector of S corresponding to the largest eigenvalue, while the columns of $\mathbf{V}$ are the other three eigenvectors. Recent work by Dr. Jared Glover have made calculating $F(\mathbf{\Lambda})$ and $\mathbf{\Lambda}$'s determination from S much more computationally efficient.

2.3 Symmetrized Bingham distribution and mixture models

The symmetrized Bingham distribution is a Bingham distribution modified to accurately reflect the symmetry of the crystals. The symmetrized Bingham distribution appears much like the Bingham distribution, except the pdf of a symmetrized Bingham distribution with paramters $\theta = \{\mathbf{\Lambda}, \mathbf{V}\}$ is the sum of Bingham distributions with parameters $\theta * \mathbf{q}_j$ across all n symmetries $\mathbf{q}_i$, or

$$p(\mathbf{g}; \theta) = \frac{1}{F(\mathbf{\Lambda})} \exp \sum_{i=1}^{4} \sum_{j=1}^{n} \lambda_i((\mathbf{v_i} * \mathbf{q_j}) * \mathbf{g}))$$

A mixture model is simply a collection of distributions with different weights, or mixing parameters. Mixture models allow for the for the representation of data using a group of simpler distributions. The probability of a quaternion orientation $\mathbf{g}$ in a symmetrized Bingham mixture with parameters $\mathbf{\Theta}$, n symmetries $\mathbf{q_j}$, and mixing probabilites α_m can be expressed as

$$p(\mathbf{g}; \mathbf{\Theta}) = \sum_{m=1}^{k} \frac{\alpha_m}{F(\mathbf{\Lambda})} \exp \sum_{i=1}^{4} \sum_{j=1}^{n} \lambda_{i,m}((\mathbf{v_{i,m}} * \mathbf{q_j}) * \mathbf{g})$$

3 Approach

3.1 Estimation maximization

Analytical solutions for fitting distributions often do not exist, and so a numerical algorithm must be employed to solve the optimization. The EM approach fits data to a mixture model without any ad-hoc parameters. The approach begins with a set n randomly initialized initial distributions. The algorithm alternates between the expectation step, or E-step, and the maximization step, or

M-step. The E-step finds the probability of each data point belonging to each of the n distributions. Normalizing these probabilities across all n distributions gives the contribution ω of that data point to each distribution, called the responsibilities. The M-step updates the parameters based on the these responsibilities. By iterating these steps, the algorithm improves the initial mixture model and allows it to accurately fit the data. The accuracy of this fit is measured by the data likelihood, which is a measure of the likelihood of the data set being generated by the current mixture model (assignments and parameters). The algorithm is terminated when the increase in likelihood is small between successive iterations. In this paper, EM was used to fit sample orientations to a symmetrized Bingham mixture model. In the E-step, the probabilities and responsibilities were calculated using Prof. Glover's computational tool. In the M-step, for each distribution, a scatter matrix $S_m = 1/N \sum_i \omega_{i,m} \mathbf{g}^{(i)} \mathbf{g}^{(i)^T} = E[\mathbf{g}\mathbf{g}^T]$ is calculated and the parameters of each symmetrized Bingham component θ_m are computed. Finally, new mixing probabilities α_m are assigned by

$$\alpha_m = \frac{\sum_{i=1}^{N} \omega_{i,m}}{\sum_{i,m=1}^{N,n} \omega_{i,m}}$$

The likelihood (L)of a sample set of orientations coming from an ODF is given by

$$L(\mathbf{G}, \mathbf{\Theta}) = \prod_{i=1}^{N} (p(\mathbf{g_i}; \mathbf{\Theta}))$$

Because a summation is preferable over a product, the log likelihood (LL) is used to measure goodness of fit, as

$$LL(\mathbf{G}, \mathbf{\Theta}) = \sum_{i=1}^{N} \log (p(\mathbf{g_i}; \mathbf{\Theta}))$$

Because the natural logarithm is strictly increasing, it still serves as a strong indicator of the likelihood. Therefore, the EM algorithm provides a way to estimate

$$\hat{\mathbf{\Theta}}_L = \arg\max_{\Theta}\{LL(\mathbf{G}|\mathbf{\Theta})\}$$

where $\hat{\mathbf{\Theta}}_L$ is the ODF with the maximum likelihood.

3.2 Minimum message length criterion

The MML criterion is used to determine the number of components in the mixture model. MML is an approach, described by Figueiredo and Jain, to prevent underfitting (too few components in the mixture) or overfitting (too many), based on the information theory idea that the simplest model which accurately describes the data is the best. Message length is the total length of information required to communicate encoded or compressed data over a communication channel. If the recipient does know the code, the message contains the econding key followed by the compressed data. The message length captures a balance between two components, the number of components in the mixture model (the key) and the data likelihood (the compressed data). The minimum ML criterion states that the best description of the data is the one that minimizes the message length, based on the information entropy of the data. Under MML, each additional mixture model component added increases the message length by a penalty related to the information entropy of the model. The increase in accuracy from the additional component must overcome this penalty. The benefit of MML is that it uses the information entropy of the model to balance between underfitting and overfitting, replacing an ad-hoc user parameter with an objective criteria. For a symmetrized

Bingham mixture model, the message length ML is expressed as

$$ML = 5\left(\sum_{m=1}^{k} \log \frac{N\alpha_m}{12}\right) + \frac{k}{2}\log\frac{N}{6} + \frac{11k}{2} - LL$$

for a mixture model of k symmetrized Bingham components with mixture probabilities α_m, N sample orientations, and log likelihood LL. By applying this criterion, the algorithm attempts to estimate

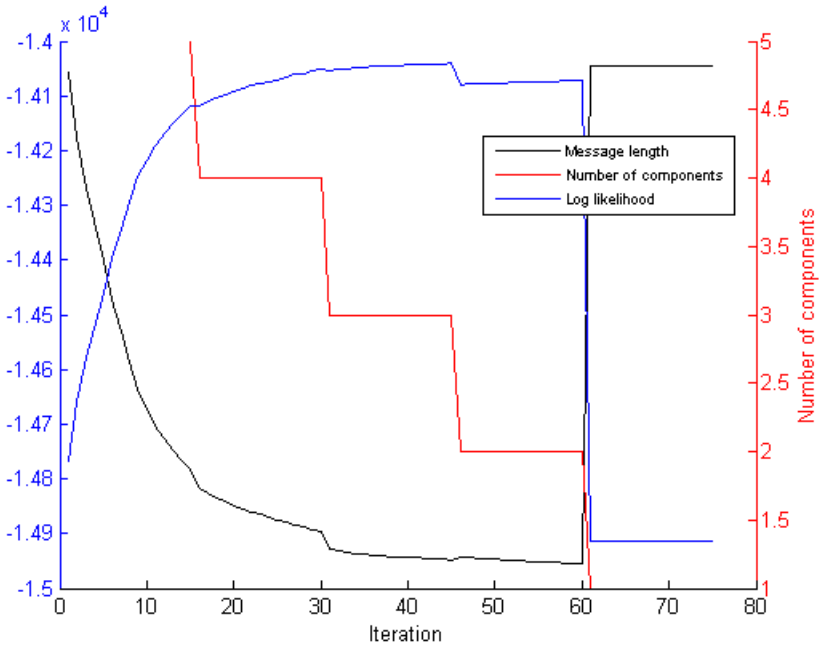
$$\hat{\boldsymbol{\Theta}}_{ML} = \arg\min_{\Theta}\{ML(\mathbf{G}|\boldsymbol{\Theta})\}$$

where $\hat{\boldsymbol{\Theta}}_{ML}$ is the mixture model with the smallest message length.

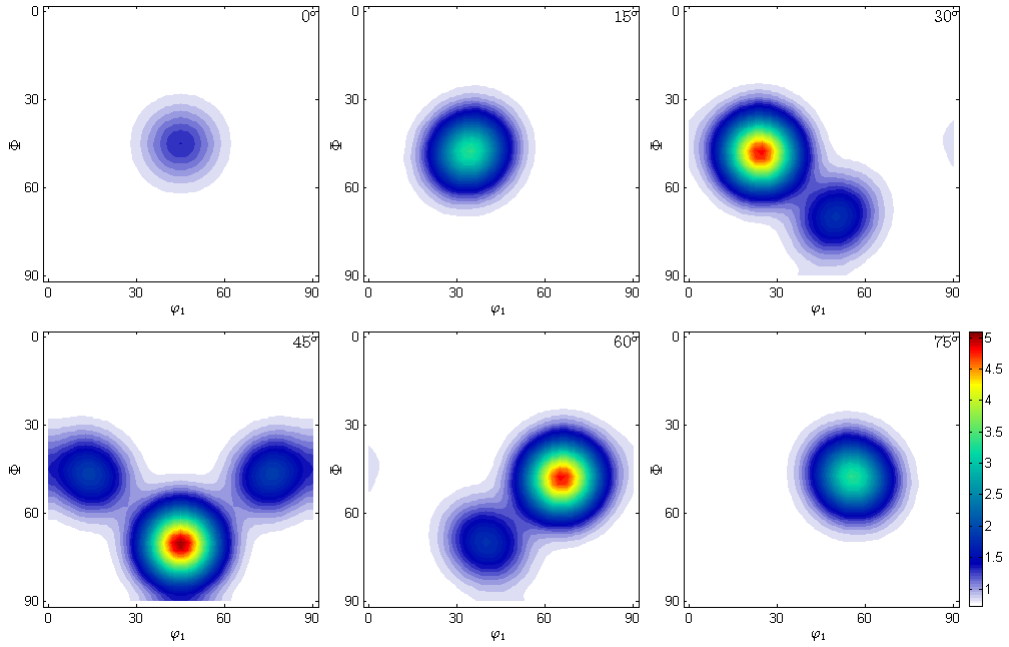
4 Results

4.1 SanteFe case study

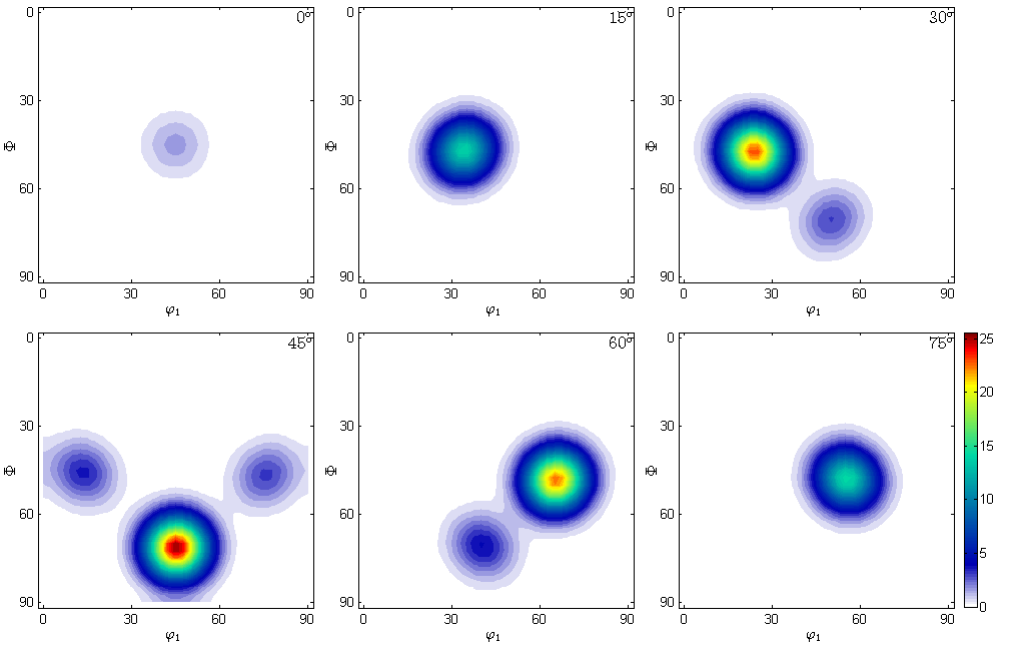
The algorithm was tested on a standard texture called the SanteFe ODF. Using a sample of 5,000 orientations drawn from this ground truth ODF, 15 EM iterations were run using 5 components. The weakest component in the mixture was removed, and the model was reevaluated for 15 more EM iterations, repeating until only one component remained. The log likelihood, message length, and number of components were recorded as a function of number of iterations.



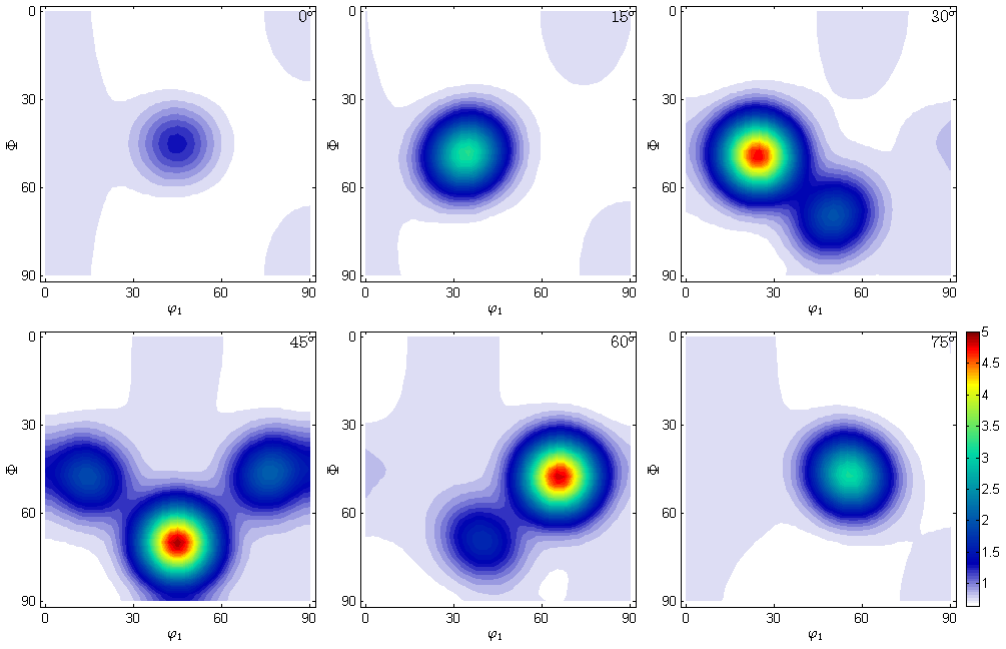
In this particular case study, although the highest log likelihood value occurred at iteration 45 with 3 components, the lowest message length value occured at iteration 60 with 2 components. By the MML criterion, the best fit was the ODF at iteration 60. To provide comparison to the MCMC method developed by Matuk, a study using the same conditions as this study was performed using that approach. His algorithm's results are included. The ground truth SantaFe ODF, the EM ODF at iteration 60, and the resulting ODF for the MCMC algorithm are shown below.



Ground Truth



EM



MCMC

While the EM algorithm properly identifies the mode, the concentration is sharply different from the ground truth, peaking at 25 in the EM ODF while peaking around 5 in the SantaFe ground truth and MCMC ODFs. The MCMC ODF uses 2 components just like the EM ODF, but provides a much better fit.

5 Conclusions

After extensive testing with different crystal symmetries, number of samples, and ground truth ODFs, the EM algorithm did not perform to expectations. Although EM algorithm will converge to a maximum likeliness, there is no guarantee that it will converge in a useful amount of time. During testing, the EM algorithm was often caught in local maxima for the likelihood. This happened in the SantaFe case study, as EM forced the ODF to become more concentrated, which increased the likelihood towards a local maximum. Another concern with the algorithm was the MML criterion.

The message length had two portions, a message length drawn from the number of components part and a log likelihood part. For samples of different size drawn from the same ODF, the message length part scaled logarithmically with the number of samples, while the log likelihood part scaled linearly with the number of samples. This meant for small samples (under 500 orientations) the MML criterion would almost never selecct an ODF with more than 2 components, while for large samples (over 10,000 samples) the number of components had almost no impact on the MML criterion. Although the MML criterion was sound in theory, in practice it impeded the ability of the algorithm to choose an ODF with better likelihood. Finally, any measure of a background uniform component (common in real samples) made convergence much more difficult, and resulted in the algorithm frequently converging towards a local maximum of the likelihood. Some adjustments, like weighting responsibilities at periodic intervals, could help the EM algorithm escape local maxima faster, but these could also impede proper convergence and would take away many of the benefits of the EM appraoch. In contrast, the MCMC algorithm showed much better performance. Because of the nature of MCMC, that algorithm can fit out of the local likelihood maxima that trap the EM algorithm. Additionally, because the MCMC method still uses symmetrized Bingham mixture models, most of the benefits of EM over the kernel density estimation approach still apply. These include avoiding human bias and permitting uncertainty quantification. MCMC's improved results come with the cost of time. MCMC requires a large number of iterations, and therefore time, to reach convergence, while the EM algorithm converges rapidly in the first few iterations. A hybridized use of the two algorithms offers some benefits. By running the EM algorithm and using the best ODF as a starting point for the MCMC algorithm, the number of MCMC iterations needed can be dramatically reduced. Likewise, locating the global ML minimum using MCMC and then using the EM algorithm to converge further permits easier uncertainty quantification. Further work

will take hybridization into account and focus on uncertainty quantification using SBD mixture models.

References

1. S.R. Niezgoda, J. Glover: *Metallurgical and Materials Transactions,* 2013, vol. 44A, p. 4891.

2. E.A. Magnuson: *Representation of Crystallographic Texture using the Symmetric Bingham Distribution*, 2016.

3. S.R. Niezgoda, E.A. Magnuson, J. Glover: *Journal of Applied Crystallography*, Submitted Februrary 2016.

4. M.A.F. Figueiredo and A.K. Jain: IEEE Trans. Pattern Anal., 2002, vol. 24 (3), p. 381.

5. C.J. Boehlert, S.C. Longanbach, T.R. Bieler: *Effect of Thermomechanical Processing on the Creep Behaviour of Udemet Alloy 188,* 2008, Philosophical Magazine.